

Case report: Identification of a Rare Homozygous Missense Variant in the PKLR gene Reported for the First Time in Transfusion-Dependent Saudi Patient

Rawabi Zahed¹, Bayan Sajer^{1*} 

¹Department of Biology, College of Science, King Abdul Aziz University, Jeddah, Saudi Arabia.

*Correspondence to: Bayan Sajer (E-mail: bsajer@kau.edu.sa)

(Submitted: 26 February 2023 – Revised version received: 21 March 2023 – Accepted: 19 April 2023 – Published online: 26 June 2023)

Abstract

Red cell pyruvate kinase deficiency is one of the most common erythrocytic glycolytic pathway defects connected with congenital non-spherocytic anemia. The condition inherited as an autosomal recessive Mendelian trait is caused by mutations in the PKLR gene located on chromosome 1q21. Pyruvate kinase enzyme is crucial in the energy-producing glycolysis pathway that provides red blood cells with the primary source of energy (ATP). We report here a case of a Saudi female patient that was initially diagnosed at a few months old with beta-thalassemia major and was treated with regular blood transfusions and iron overload management. At the time of our sample collection, the patient was recently transferred to King Abdul Aziz University Hospital. Genetic testing was performed to identify the disease-contributing variant of beta-thalassemia using TaqMan genotyping of six common beta-thalassemia variants (negative results). NGS targeted HBB gene sequencing which did not reveal any related variants. MLPA was performed to rule out alpha thalassemia diagnosis. The use of whole genome sequencing revealed a rare missense variant in the PKLR gene c.1015G > A (D339N) in a homozygous state that correlates to her severe phenotype. Documenting this incident will aid medical staff in providing appropriate care to similar cases and highlights the importance of following up with the diagnosis investigation process to minimize misdiagnosis incidences.

Keywords: PKD; heterogeneity; β -thalassemia; consanguineous family; PKLR gene

Introduction

The work described in this section contributed to the manuscript “Identification of a rare homozygous missense variant in the PKLR gene reported for the first time in transfusion-dependent Saudi patient” submitted to *Frontiers in Genetics*, section Genetics of Common and Rare Diseases (manuscript ID: 1221722, submitted on: 12 May 2023).

The most common erythrocytic glycolytic pathway defect connected with congenital non-spherocytic anemia is red cell pyruvate kinase deficiency. The condition which is inherited as an autosomal recessive Mendelian trait is caused by mutations in the PKLR gene located on chromosome 1q21. Our knowledge of the genetic diversity, pathophysiology, and hemolytic anemia complications brought on by red cell pyruvate kinase deficiency has grown over the past few decades. More than 300 distinct variants have been discovered since the discovery of the first PK deficiency pathogenic variants in 1991, and extensive research has been done on the investigation of its molecular mechanisms and the existence of genotype-phenotype connections.¹

Pyruvate kinase deficiency (PKD) was initially identified in the early 1960s. De Gruchy and colleagues reported that adenosine triphosphate ATP, but not glucose, was able to

reverse the symptoms of non-spherocytic hemolytic anemia in a subset of patients by performing autohemolytic test on patient's red blood cells alone at first, then compares it with autohemolytic rate of blood with added glucose, blood with added adenosine, and blood with added adenosine triphosphate using modified Dacie's method. This discovery came after Selwyn and Dacie's initial discovery of the connection between glycolysis and hemolytic anemia in the 1950s. PKD was soon identified as the molecular cause of this anemia by Valentine and Tanaka and their colleagues.²⁻⁶

Pyruvate kinase (PK) is crucial in the energy-producing glycolysis pathway (the Embden-Meyerhoff pathway) that provides the red blood cell with the primary source of energy (ATP) by breaking down the sugar glucose. PK catalyzes and synthesizes the conversion of phosphoenolpyruvate to pyruvate and one molecule of ATP.⁷

The *PKLR* gene, which is 9.5 kb in size and has 12 exons, is found on chromosome 1q21. In accordance with tissue-specific promoters, the gene encodes for the enzyme's liver (L) and erythrocyte (R) isoforms. Both the erythrocyte and the hepatic isoenzyme share 10 of the 12 exons that make up the coding area, whereas exons 1 and 2 are unique to each isoform. The 2060 bp long cDNA for *PKR* codes for 574 amino acids.¹

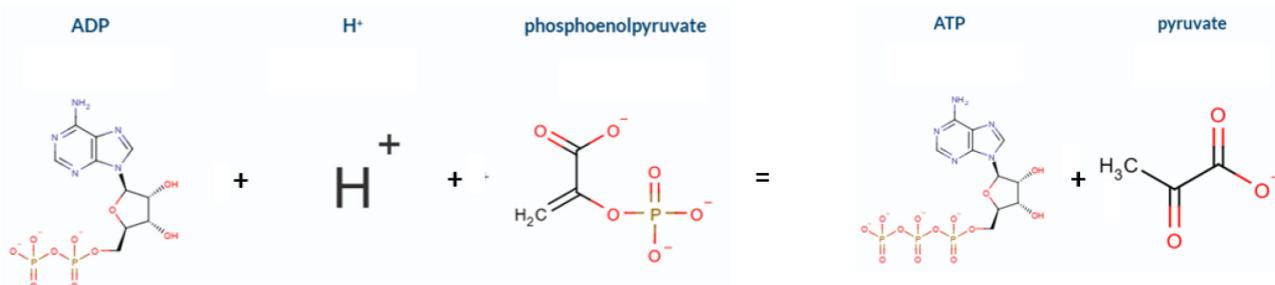


Fig. 1 Equation of pyruvate kinase catalytic reaction.

The PKLR gene, which is active in both the liver and red blood cells, codes for both pyruvate kinase isoenzymes PK-L and PK-R. Since mature RBC lack mitochondria and have a lifespan of 100 to 120 days, they rely on ATP for preserving their structural and functional integrity. Mutations in the PKLR gene result in abnormal or deficient PK enzyme which leads to insufficient ATP synthesis, loss of RBC membrane flexibility, cellular dehydration (As a result of the exhaustion of the energy supply for membrane ion transporters), and early destruction of RBCs in the spleen or liver causing hemolytic anemia and enlarged spleen. The hemoglobin-oxygen dissociation curve shifts rightward due to the accumulation of 2,3-bisphosphoglycerate when PK activity is inadequate. Shortness of breath, weariness, and pallor (pale skin) are symptoms of an insufficient amount of oxygen transporter RBCs.⁸ The severity of the resulting congenital non-spherocytic anemia varies from patient to patient. When RBCs are prematurely destroyed, iron and a chemical known as bilirubin are released, which causes an excess of these substances to circulate in the blood. Jaundice (yellowing of the eyes and skin) is a result of elevated bilirubin levels in the blood. If these levels went untreated, these levels could accumulate and have serious effect that can cause brain damage (acute bilirubin encephalopathy) in infants (very rare to develop in adults). Also, high levels of bilirubin make gallstones more likely to form.⁹

The diagnosis, treatment, and follow-up of children and adults with PK deficiency lack standards of care at present. Because it is unclear how people with PK deficiency differ in terms of risk of problems or clinical course, management techniques frequently mirror those used for patients with hereditary spherocytosis or thalassemia intermedia.¹⁰

Worldwide reports of PKD patients have been made since the condition was first described. It is found in all ethnic groups. Although the frequency of PKD is not known with certainty, it is estimated to occur in 3 to 8 out of every 100,000 people.¹¹ Cases may be misdiagnosed due to its rarity, difficulties in making a diagnosis, and the wide range of clinical symptoms; as a result, this frequency might be understated. Because heterozygous carriers frequently exhibit no symptoms, it is challenging to estimate the prevalence.

The number of PKLR gene pathogenic variants that are now known is constantly growing. Canu and colleagues have reported 260 mutations.¹² On the other hand, HGMD has reported 290 pathogenic variants in their database. The coding regions are where the vast majority of pathogenic mutations are found, with most of them being missense mutations distributed throughout all PKLR's exons. The majority of PK deficiency variations impact residues important to the enzyme's structure and/or function. The wide range of clinical manifestations associated with PKD reflect the significant molecular heterogeneity, and the pursuit of a link between genotype and phenotype has been the subject of research for many years.

Diagnosis

Hematologically, PK deficiency presents like any other hereditary hemolytic illness. Clinical presentation and test indicators of chronic hemolytic anemia, such as severe jaundice, splenomegaly, an increase in reticulocytes, and hyperferritinemia, might raise the possibility of PKD as a diagnosis. With the exception of miscarriages and affected siblings, the family history of PKD is largely unrevealing because it is an

autosomal recessive illness. The diverse range of inherited and acquired hemolytic diseases is part of the differential diagnosis. The diagnosis is made when more frequent causes of hemolysis have been ruled out, reduced PK enzyme activity has been shown (enzyme testing). Although low enzyme activity is correlated with PKD diagnosis, normal enzyme levels can be falsely diagnosed as normal in patients who have received blood transfusions because of contamination with normal donor cells. Additionally, transfusion dependency without obvious etiology since birth, with unexplained hyperbilirubinemia are strong indicatives of the disease. Molecular analysis and genetic diagnosis of PKD is considered the best way to confirm the disease if compound heterozygous or homozygous pathogenic mutations in the PKLR gene have been found.¹³

Case Presentation

Peripheral blood was collected from a female patient in her early 20s at the day care center of King Abdulaziz University Hospital (KAUH) where she receives her monthly blood transfusion. The patient was previously diagnosed as suspected beta-thalassemic with symptoms presented from 3–4 months after her birth and was administered regular continuous transfusions accordingly. Sample DNA was investigated for beta-thalassemia variations using several molecular techniques, including TaqMan genotyping and NGS targeted sequencing of *HBB* gene using Ion Torrent PGM. Variations on *HBB* gene were never identified to support the suspected diagnosis. Additionally, alpha thalassemia was investigated using MLPA technique but yielded negative results. Further investigation of the case and reviewing clinical data and family history became necessary to correlate her genotype with the severity of her phenotype.

Clinical History

When the proband was born, she was discharged as a healthy new-born. Within a few days or weeks, the baby's skin and whites of her eyes started to appear yellowish with the family unaware of the severity of their daughter's condition, they did not seek medical attention until around 3–4 months old. Jaundice (yellow skin and eyes) is a condition in which the infant's liver may not be developed enough to effectively eliminate bilirubin which leads to accumulation of this substance in the infant's body. Kernicterus is a disorder that develops when severe jaundice is left untreated for an extended period of time and can cause brain defects, hearing impairment, and athetoid cerebral palsy. Unfortunately, when the family returned to seek medical care, brain damage and hearing impairment was already present in the proband had severe hemolytic anemia of extremely low RBCs count and low hemoglobin levels that required immediate blood transfusion. At that time, a diagnosis of suspected beta-thalassemia was given to her to fast-track the approval of her transfusion and regular monthly appointments were set on her behalf for continuous transfusions. The patient was not investigated further until she was transferred to be treated at KAUH (around the time of sample collection for our study). At that time, after sample collection and molecular characterization of the *HBB* gene, the hospital was informed that there was no genetic finding that correlates

with beta-thalassemia. An attending hematologist reviewing her case file decided to investigate her symptoms and family history and performed Hb electrophoresis for the proband's mother (father deceased) and her healthy siblings and found no evidence of beta-thalassemia. The patient is the offspring of consanguineous parents and because of the rarity of pathogenic alleles in the population, it would be extremely rare to find cases in non-consanguineous families. Because it is a recessive disorder, there is unlikely to be any phenotype in the parents or even heterozygous siblings. There was no history of transfusion dependency in the family.

An interview with the proband's caring sister was conducted by phone to collect information on the family history of the disease, as well as the disease history of the proband. No history of the disease was detected in any of the proband's immediate or secondary family members. The parents were consanguineous but there was no evidence of the disease in other family members. Additionally, the attending hematologist was also contacted to further understand the proband's clinical manifestation and help correlates her phenotype with her genotype.

To identify a different genetic etiology, whole exome sequencing was performed rather than HBB-directed sequencing. NovaSeq 6000 DNA Exome by Illumina was Cite Table 1. Procedure was carried out as mentioned in the manufacturer's user guide (explained in detail in Experimental Procedure chapter). Software used for data analysis was Bcl2fastq v2-20, BWA-mem (aligner) v-0.7.12, Samtool v1.2, GATK-HaplotypeCaller v-4.1.4.

Results

After molecular investigation using various sequencing techniques, neither TaqMan genotyping nor NGS targeted sequencing revealed any mutation related to her initial suspected diagnosis of beta-thalassemia, nor any deletions in her alpha globin genes related to alpha thalassemia.

A score ratio of 1 indicates no change in copy number compared to reference sample. On the other hand, a score of 0.5 indicates a heterozygous deletion while a score of 1.5 indicates a heterozygous duplication. The blue boxes represent the 95% confidence interval of a probe over the reference samples. The colored dots are the calculated probe ratio. The error bars

surrounding the dots represent a 95% confidence interval estimate for each probe in a sample. The red and blue lines display lower arbitrary border and upper arbitrary border respectively. These borders are ± 0.3 from the average value of a probe over the reference sample. The black dot indicates no change in probe copy number because the 95% confidence interval estimates (the error bars) of the probe overlap with the 95% confidence intervals of the same probes over the reference sample. The purple dot indicates a decreased signals of more than two standard deviations compared to the reference samples, but the lower arbitrary border has not been crossed. The pink dot indicates an increased signals of more than two standard deviations compared to the reference samples, but the upper arbitrary border has not been crossed. The red dots indicate a decreased signals of more than two standard deviations compared to the reference samples and the lower arbitrary border has been crossed. The red dot far down in the ratio chart with one error bar indicates no signal found. Source: Coffalyser.Net reference manual v01.

WES analysis was then used to get a detailed view of the protein-coding regions of her whole genome. we have generated a gene list to include genes that have been reported to be involved in one way or another in hereditary anemia disorders. Generating gene panels to investigate inherited hemolytic anemia disorder have been previously employed.^{14,15} Our list was assembled by looking into literature and several gene panels that are used by clinical laboratories to test for anemia related disorders. This list included, ABCB7, ABCG5, ABCG8, ADA, ADA2, ADAMTS13, ADH5, AK1, ALAS2, ALDOA, AMN, ANK, ANK1, ATRX, BCL11A, BLM, BRCA2, BPGM, BRIP1, C15ORF41, CD59, CDAN1, CDIN1, CYB5A, CYB5R3, COL4A1, CUBN, DKC1, DHFR, DNAJC21, DNASE2, EFL1, EBP41, EBP42, ENO1, ERCC4, ERFE, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCM, FANCL, FTCD, G6PD, GATA1, GCLC, GIF, GLUT1, GLRX5, GPC, GPI, GSR, GPX1, GSS, GYPC, HEATR3, HK1, HSCB, HSPA9, HMOX1, KCNN4, KIF23, KLF1, LARS2, LPIN2, MTR, MTRR, NDUFB11, NHLRC2, NHP2, NT5C3A, NRF1, PALB2, PC, PDHA1, PDHX, PIEZO1, PFK, PFKM, PGK1, PKLR, PUS1, RAD51C, REN, RHAG, RPL11, RPL15, RPL18, RPL27, RPL31, RPL35A, RPL5, RPS10, RPS17, RPS19, RPS24, RPS26, RPS27, RPS28, RPS29, RPS7, SBDS, SEC23B, SLC19A1, SLC11A2, SLC19A2, SLC25A38, SLC4A1, SLC2A1, SF3B, SLX4, STEAP3, SPTA1, SPTB, SRP54, TCN2, TF, THBD, TMPRSS6, TRNT1, TSR2,

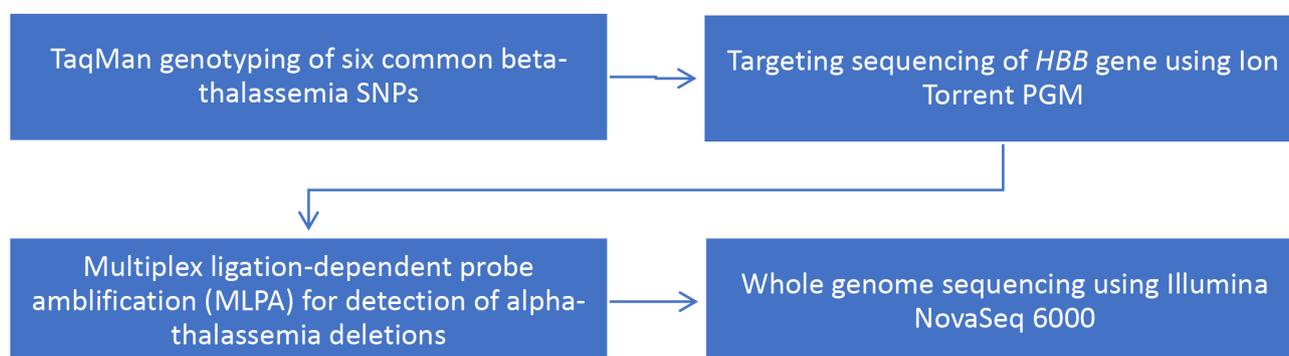


Fig. 2 Visual diagram of molecular techniques utilized in attempt to identify disease causing variations of sample 183. First, TaqMan genotyping was performed on DNA sample using six of the most frequent beta thalassemia variations in Saudi Arabia. With negative results from the genotyping, targeted sequencing of HBB gene using Ion Torrent PGM was performed but no variation was detected. After that, to rule out alpha thalassemia, we performed MLPA testing, and no deletions were detected. Finally, we detected the PKLR variant using whole genome sequencing.

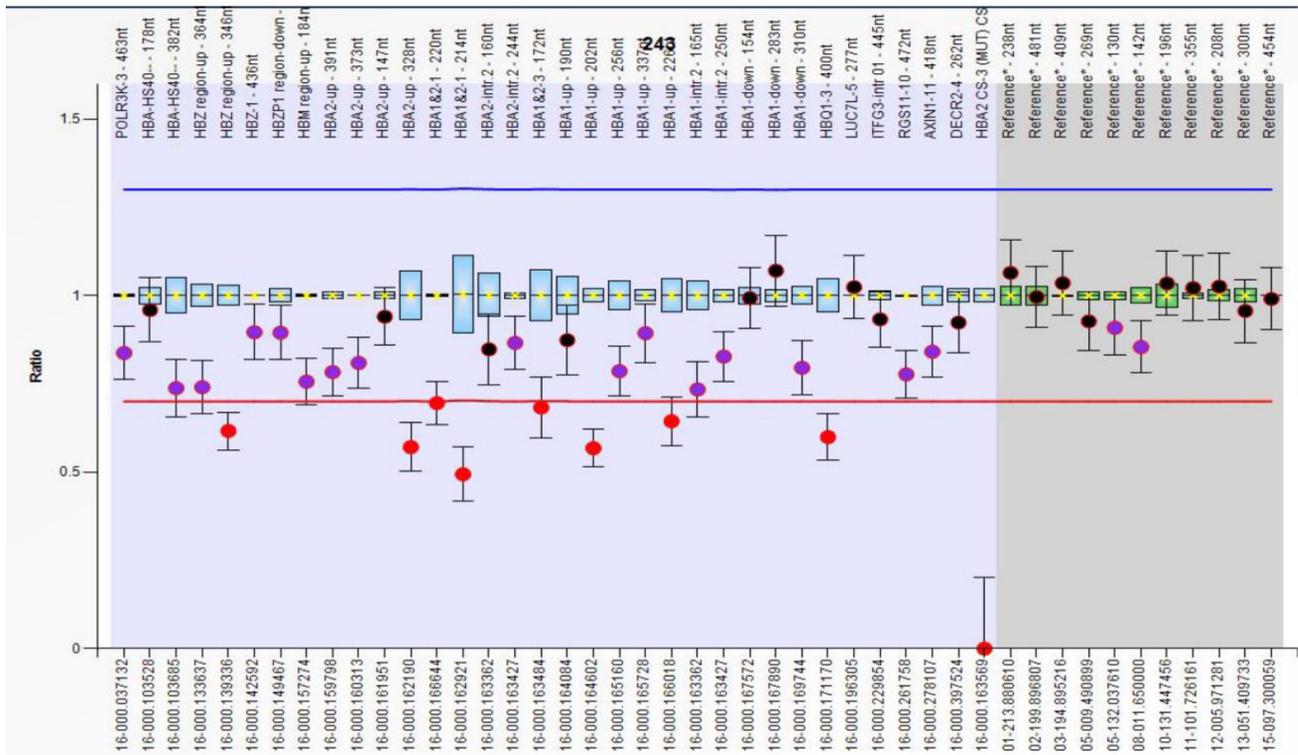


Fig. 3 MLPA reaction result showing no significant deletions.

Table 1. Whole exome sequencing result of patient 183.

Sample	HGVS name	Mutation Status	Genomic Location on Assembly GRCh37	Location on gene	Molecular consequence	Reference cluster ID (rs#)
14-0183	NM_000298.6 c.1015G>A	Homozygous	Chr1: 155264127	Exon 7	Missense SNV NP_000289.1 P.Asp339Asn	rs747097560 C-T

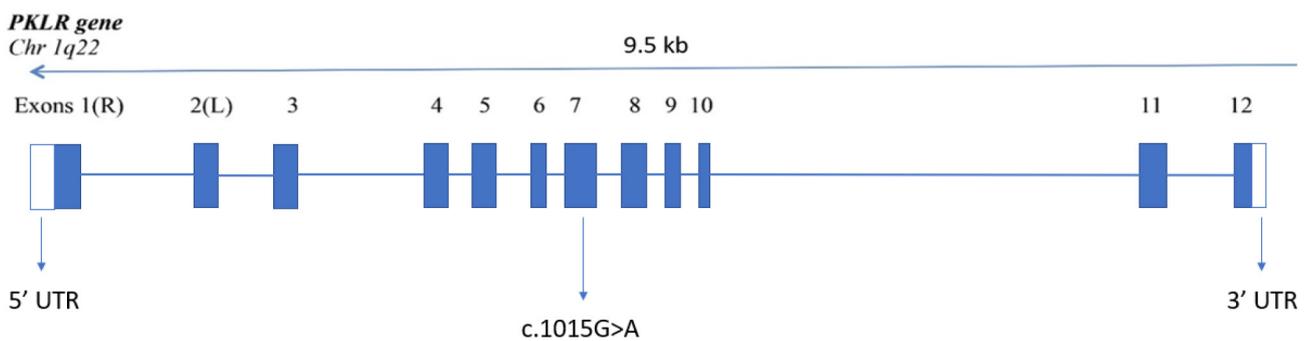


Fig. 4 c.1015G>A variation mapped on PKLR gene. R indicates erythrocytic isozyeme, while L indicates hepatic isozyeme which are both encoded by the same gene but with different promoters. With 574 amino acids, the R-type subunit is 31 residues longer than the L-type at its amino terminus.

TPI1, UMPS, UGT1A1, VPS4A, XK, YARS2. The advantages of using gene panels as a filtration search scenario include shorter turnaround times, easier data processing, more coverage in the regions of interest, and fewer incidental results.¹⁶⁻¹⁸

A homozygous missense mutation was found on exon 7 of PKLR gene that had high potential to correlate and explain her symptoms and phenotype. This missense mutation leads to a change of amino acid number 339 both forms translated protein from aspartic acid to asparagine (D339N). The overall allele frequency of this variation in the gnomAD v2.1.1 dataset

is 0.002%. This variation has not been reported before in Saudi population.

The hematologist was presented with clinical manifestation supporting of PKD that includes hemolytic anemia with no known etiology, and severe neonatal jaundice that progressed to bilirubin encephalopathy with lifelong neurological damage, the hematologist investigated PKD as potential diagnosis which was clinically confirmed by testing pyruvate kinase enzyme activity levels (before transfusion) which appeared to be low.

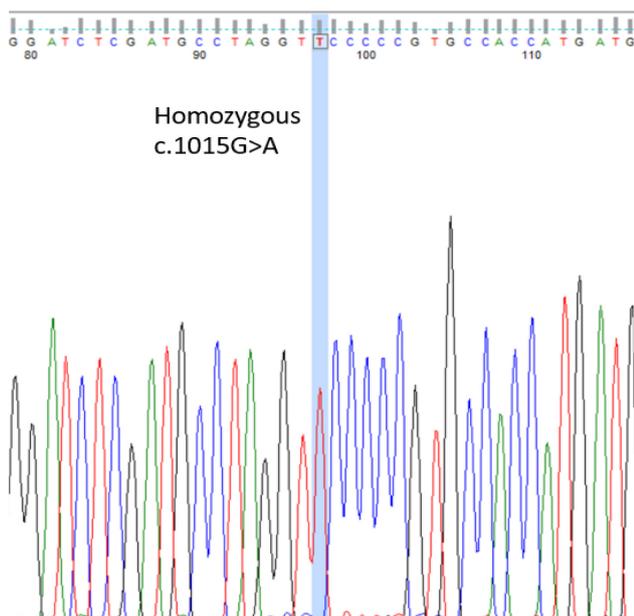


Fig. 5 Sanger validation of PKLR mutation.

Score	Expect	Identities	Gaps	Strand
329 bits(178)	8e-86	180/181(99%)	0/181(0%)	Plus/Plus
Query 1	GTCGCTATTCCTCCATCACCTTTCTTCCTGCTGCCTCTGCTTGTATTCCTCCAACTTC	60		
Sbjct 7496	GTCGCTATTCCTCCATCACCTTTCTTCCTGCTGCCTCTGCTTGTATTCCTCCAACTTC	7555		
Query 61	TCAGGTTTGTAGAAATCCTGGAGGTGAGCGACGGCATCATGGTGGCACGGGGAACTAG	120		
Sbjct 7556	TCAGGTTTGTAGAAATCCTGGAGGTGAGCGACGGCATCATGGTGGCACGGGGAACTAG	7615		
Query 121	GCATCGAGATCCCAAGCAGAGAAAGTTTTCTTGGCTCAGAAGATGATGATTGGCGCTGCA	180		
Sbjct 7616	GCATCGAGATCCCAAGCAGAGAAAGTTTTCTTGGCTCAGAAGATGATGATTGGCGCTGCA	7675		
Query 181	A 181			
Sbjct 7676	A 7676			

Fig. 6 Alignment of PKLR gene Sanger Sequence.

Data Analysis

The homozygous missense variant found in our proband had not been associated with PKD phenotype except for a recent case report in 2022. Rehman and his group have reported this bi-allelic variant in the PKLR gene and correlated them to PK deficiency found in a consanguineous Pakistani family¹⁹ in which there were four affected children.

The PKLR gene encodes for the pyruvate kinase enzyme, which catalyzes the transfer of a phosphate group from 2-phosphoenolpyruvate (PEP) to ADP, forming pyruvate and ATP. This is the final stage in the glycolytic process, and under normal conditions, the reaction is not reversed. Mammals, including humans, possess four distinct PK isozymes. Alternative splicing of PKM gene transcripts results in mRNAs encoding either the M1 (muscle) or M2 (fetal) proteins. The sole difference between PKM1 and PKM2 mRNAs is whether they include one of two overlapping exons or not, PKM1 (8-9-11) while PKM2 (8-10-11).²⁰ Fructose 1,6-bisphosphate (FBP) and phosphoenolpyruvate (PEP) allosterically activate M2 PK, while simple saturation hyperbolic kinetics are exclusive to the M1 enzyme. Using tissue-specific alternative promoters, the same PKLR gene codes for the mammalian liver and erythrocyte PK isozymes. PEP and FBP stimulate isozymes in both erythrocytes and the liver.^{21,22}

PK enzyme is a tetramer made up of four subunits. Each subunit has four domains: A domain with (β/α) 8 barrel

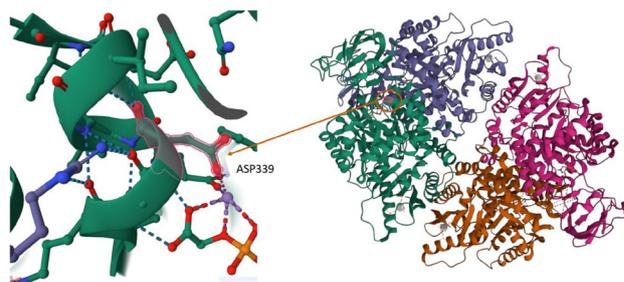


Fig. 7 Pyruvate Kinase 3D structure adapted from PDB database (2VGI).²³

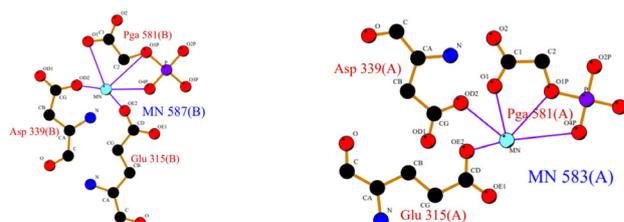


Fig. 8 Crystallography Ligand with 2-phosphoglycolic acid/metal interaction involving metal ion Mn²⁺ in both chain A and chain B using LIGPLOT²⁶ to Asn would remove the negative charge from the ligand.

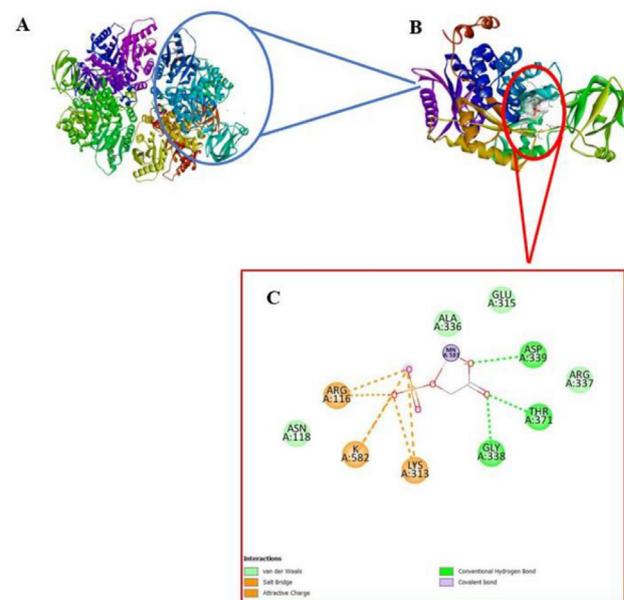


Fig. 9 Pyruvate Kinase active site. A represents the 3D structure of the enzyme tetramer. B is a single chain in the enzyme. C represents the 2D structure of the enzyme's active site bound to its substrate 2-phosphoglycolic acid. The homology modelled 3D structure of pyruvate kinase using I-TASSER server.

topology called domain A, domain B which is located between alpha-helix 3 and beta-sheet 3 of domain A, domain C with α+β topology, and a small helical domain called the N-terminus domain.²³ The control of PK activity relies heavily on this multidomain structure. The activation of the enzyme involves a series of rotations in the enzyme's domains and subunits, as well as changes to the geometry of the enzyme's active site. Residues at the domain and subunit interfaces play

critical roles in this mechanism by mediating interactions between the activator-binding site (inside the C domain) and the catalytic core (between the domains A and B).^{1,24} Fructose 1,6-bisphosphate acts as an allosteric activator of red cell pyruvate kinase, whose reaction kinetics follow a sigmoidal shape with respect to PEP.

The catalytic mechanism of pyruvate kinase involves metal binding, a Mg²⁺ and K⁺ ions, that is coordinated by several residues in the active site. The ions play a critical role in stabilizing the negative charges that are found during the reaction, facilitating the transfer of the phosphate group. The metal binding properties of pyruvate kinase are important for understanding the enzyme's function and regulation.²⁵

The carboxylate side chain of Asp339 is involved in the coordination of the Mg²⁺, which is critical for the enzyme's catalytic activity. In studies of X-ray crystal structure, the inhibitor 2-Phosphoglycolic acid is bound at the PEP-binding site of the A domain, where protein residues and the cations Mn²⁺ and K⁺ form a complex network of hydrogen bonds. K⁺ ions and the side chain of Arg116 bind the phosphate group, whereas Mn²⁺ ions, the side chain of Thr371, and the main chain nitrogen atoms of Gly338 and Asp339 at the N terminus of a short helical segment belonging to loop 6 of the A domain, anchor the carboxylate moiety. These interactions help to stabilize the Mg²⁺ ion and orient it in the active site, allowing it to facilitate the transfer of the phosphate group from PEP to

ADP.²² Mutations in Asp339 which involves Mg²⁺ binding can impair the catalytic activity of pyruvate kinase and alter the substrate specificity or allosteric regulation of the enzyme, leading to changes in metabolic pathways and cellular functions.

Asp339 is a highly conserved residue located in the active site of pyruvate kinase. Based on a search of the UniProt database, Asp339 is highly conserved across different organisms, suggesting its functional importance in the enzyme. Asp33 is conserved in all vertebrate pyruvate kinase isoforms, including human (P30613), mouse (P52480), zebrafish (Q9I9N4), Xenopus (Q90XJ9), and is also conserved in other organisms such as fruitfly (P07132) and yeast (P00887).

Aspartic acid (Asp) and asparagine (Asn) are structurally similar amino acids, with Asn having an additional amide group in its side chain. However, Asp is negatively charged, while Asn is uncharged. Therefore, the substitution of Asp339 with Asn would eliminate the negatively charged carboxylate group that is crucial for the coordination of Mg²⁺ in the active site. This could lead to a reduction or even loss of Mg²⁺ modelled binding, impairing the catalytic activity of pyruvate kinase, potentially leading to significantly functional consequences, metabolic disorders, and physiological effects affecting its interaction with other molecules and regulatory factors.

On a similar note, in 1997, a variation that produces the amino acidic substitution Asp390Asn was found by Zanella and his research group while studying a group of Italian patients with pyruvate kinase deficiency.²⁹ At the A/A interface, Arg337 and Ser389 from two separate subunits form a hydrogen bond network that prevents Asp390 from interacting with the solvent. Crystallographic analyses of *E. coli* PK reveal that Asp390, via connecting quaternary structural alterations with active site rearrangements, is essential for the enzyme's allosteric transition. Molecular investigation reveals that replacing Asp390 with Asn results in a nearly inactive protein that is, nonetheless, just as thermostable as wild-type RPK. This suggests that the Asp390Asn mutation prevents the protein from transitioning to its active R state.³⁰⁻³²

In our investigation of the variation, we have modelled of the mutated protein and compared it with the normal protein, which resulted in noticeable changes with regards to their protein-ligand 2D chemical interactions, hydrophobicity, H-bonds, and charges.

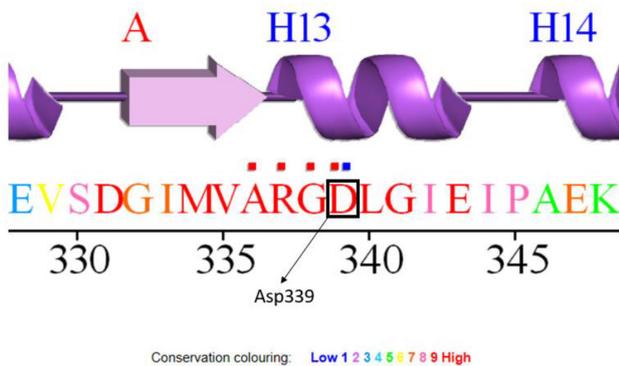


Fig. 10 Amino Acid sequence of PKLR gene showing Asp339 residue conservation. Residue conservation calculated by ConSurf-DB²⁷

277	RRFD---EILEASDGMVARGDLGIEI PAEKVFLAQKMMIGRCNRAGKPVICATQMLESMIKKPRPTRAEGSDVANAVLD	353	cat
248	NNFD---EILKVTGDMVARGDLGIEI PAPEVLA VQKKLI AKSNLAGKPVICATQMLESMTYNPRPTRA EVDVGNAILD	324	baker's yeast
228	NNFD---EILEASDGMVARGDLGVEIPVEEVI FAQKMMIEKCI RARKVVITATMMLDSMIKNNRPTRAEAGDVANAILD	304	Escherichia coli
262	NNFD---EILEETDGMVARGDLGIEI PAQKVFIAQKMMIAKCN IKGKPVICATQMLESMTYNPRPTRA EVDVGNAVLD	338	Emericella nidulans
271	NNFA---EILEETDGMVARGDLGIEI PAAEVFAAQKMMIAMCNIAGKPVICATQMLESMIKNNRPTRAEISDVGNAVTD	347	Hypocrea jecorina
259	NNFD---EILKETDGMVARGDLGIEI PAPQVFI AQKQLIAKCNLAGKPVICATQMLDSMTYNPRPTRA EVDVGNAVLD	335	Yarrowia lipolytica
248	INFN---EILRETDSFMVARGDLGMEIPVEKIFLAQKMMIYKCNLAGKAVVITATQMLESMIKSPAPTRA EATDVANAVLD	324	potato
229	ANID---EILEAADGLMVARGDLGVEI PAEEVPLIQKLLIKKCNMLGKPVITATQMLDSMQNRPRTRAEASDVANAI FID	305	Geobacillus stearo...
231	ANDEamdDII LASDVMVARGDLGVEI GPELVGVQKKLIRRSRQLNRAVITATQMMESMISNMPPTRAEVM DVANAVLD	310	Haemophilus influe...
229	DNID---EILQVSDGLMVARGDMGVEI PFINVPFVQKTLIKKCNALGKPVITATQMLDSMQENRPTRAE VTDVANAVLD	305	Lactobacillus delb...
248	NNFD---EILKVTGDMVARGDLGIEI PAPEVLA VQKKLI AKSNLAGKPVICATQMLESMTYNPRPTRA EVDVGNAILD	324	baker's yeast
228	NNFD---EILEASDGMVARGDLGVEI PVEEVI FAQKMMIEKCI RARKVVITATMMLDSMIKNNRPTRAEAGDVANAILD	304	Escherichia coli
262	NNFD---EILEETDGMVARGDLGIEI PAQKVFIAQKMMIAKCN IKGKPVICATQMLESMTYNPRPTRA EVDVGNAVLD	338	Emericella nidulans
271	NNFA---EILEETDGMVARGDLGIEI PAAEVFAAQKMMIAMCNIAGKPVICATQMLESMIKNNRPTRAEISDVGNAVTD	347	Hypocrea jecorina
259	NNFD---EILKETDGMVARGDLGIEI PAPQVFI AQKQLIAKCNLAGKPVICATQMLDSMTYNPRPTRA EVDVGNAVLD	335	Yarrowia lipolytica
248	INFN---EILRETDSFMVARGDLGMEI PVEKIFLAQKMMIYKCNLAGKAVVITATQMLESMIKSPAPTRA EATDVANAVLD	324	potato
229	ANID---EILEAADGLMVARGDLGVEI PAEEVPLIQKLLIKKCNMLGKPVITATQMLDSMQNRPRTRAEASDVANAI FID	305	Geobacillus stearo...
231	ANDEamdDII LASDVMVARGDLGVEI GPELVGVQKKLIRRSRQLNRAVITATQMMESMISNMPPTRAEVM DVANAVLD	310	Haemophilus influe...
229	DNID---EILQVSDGLMVARGDMGVEI PFINVPFVQKTLIKKCNALGKPVITATQMLDSMQENRPTRAE VTDVANAVLD	305	Lactobacillus delb...

Fig. 11 The amino acid, D (Asp339), conservation across different organisms. Source: CCD database (cd00288).²⁸

Molecular docking was performed for the normal PKLR protein and the mutated PKLR (Asp339Asn) domains. First, we used both NCBI (PDB ID: 2VGB) and UniProt (ID: P30613) databases for the biological data collection of normal PKLR enzyme. I-TASSER server was used for modeling of the 3D protein structures of normal and mutated PKLR enzyme, and each compound structure was obtained

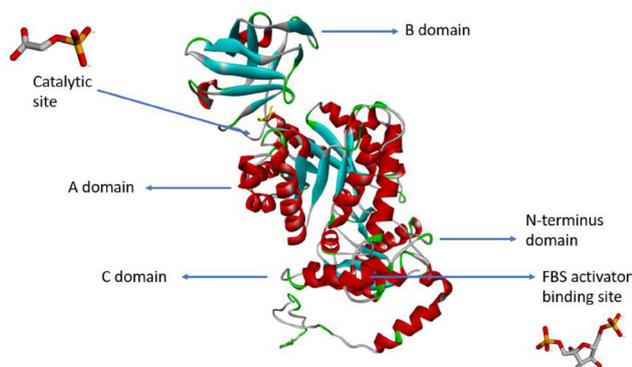


Fig. 12 **Pyruvate kinase 3D structure. The homology modelled 3D structure of rs747097960 variation using I-TASSER server.**

from PubChem database. For energy minimizing process, Swiss PDB Viewer (spdbv) was used to modify all compound structures. Format conversion from pdb to pdbqt was done by using Open Babel (Version 2.3.1) software. After that, Discovery Studio software (Version 2019) was used to handle the alteration of each protein and ligand by adding hydrogen atoms and metals, before the molecular docking process. Auto Dock Vina (Version 2.0) was used to define the grid box with 1.00 Å spacing and a grid map of 72 X, 76 Y, 84 Z Å points for the mutated PKLR with the predicted new substrate (proline) and 14 X, 16 Y, 10 Z Å points for the mutated PKLR with its normal substrate. Auto Dock was used to rank Van der Waals interactions, binding energy and inhibition constant.³³

The analysis of the binding sites revealed the pockets for each protein and the nature of the pockets using the Discovery studio program. The protein pockets and the amino acids involved in binding can help us make an estimate of how H-bonds, chemical interactions, hydrophobic/hydrophilic, positively/negatively charged, and what type of substrate may bind to that specific protein.

The 2D chemical interactions comparison between the normal and mutated PKLR showed that the mutated form will

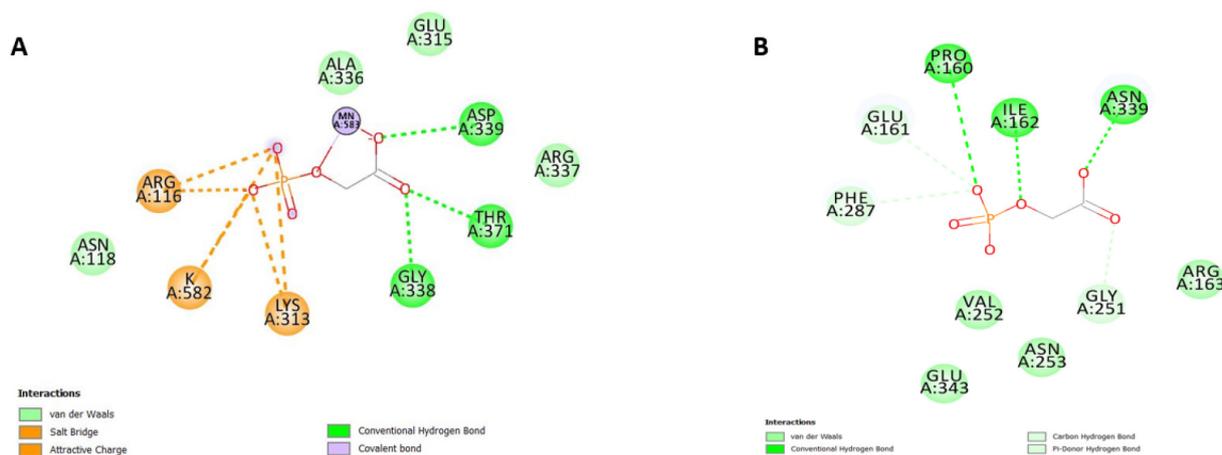


Fig. 13 **A comparison in 2D structure of the active site chemical interactions. A) The normal domain of PKLR. B) The mutated domain with "2-phosphoglycolic acid".**

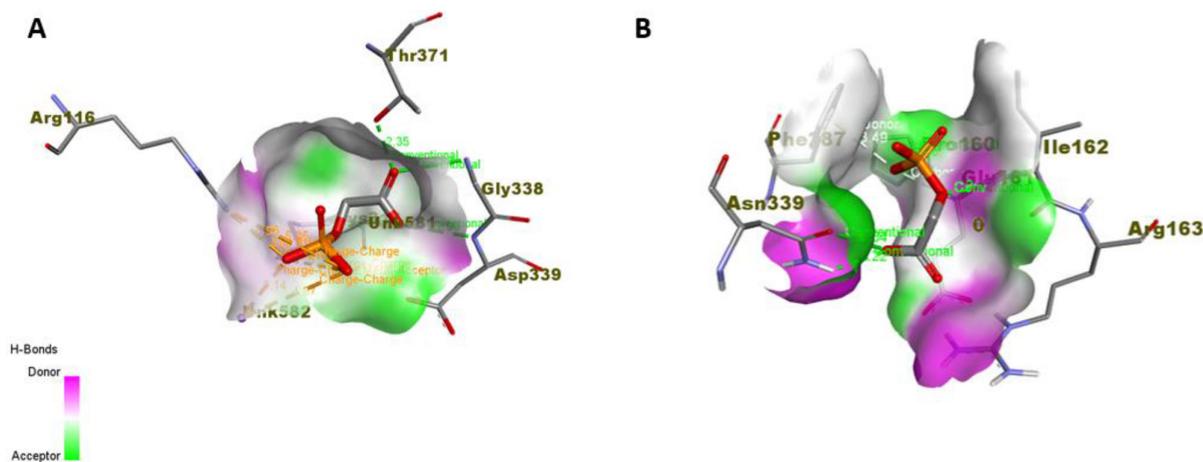


Fig. 14 **A comparison in H-bonds. A) The normal domain of PKLR. B) The mutated domain with "2-phosphoglycolic acid". H-bonds as donors appeared with pink color, while H-bonds as acceptors appeared with green color. H-bonds as donors through the docked of both mutated PKLR complex were more than the normal chain-substrate complex.**

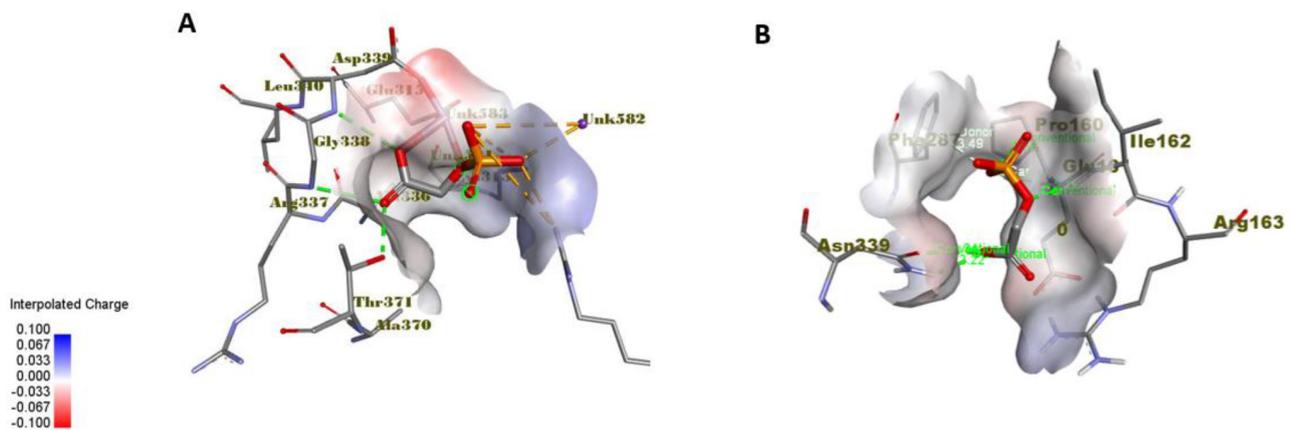


Fig. 15 A comparison of protein's charges.

A) The normal domain of PKLR. B) The mutated domain with "2-phosphoglycolic acid". Positive charges appeared with blue color while negative charges appeared with red color. The docked chain A of normal PKLR both negative and positive charges, while the docked mutated PKLR charges were mostly neutral.

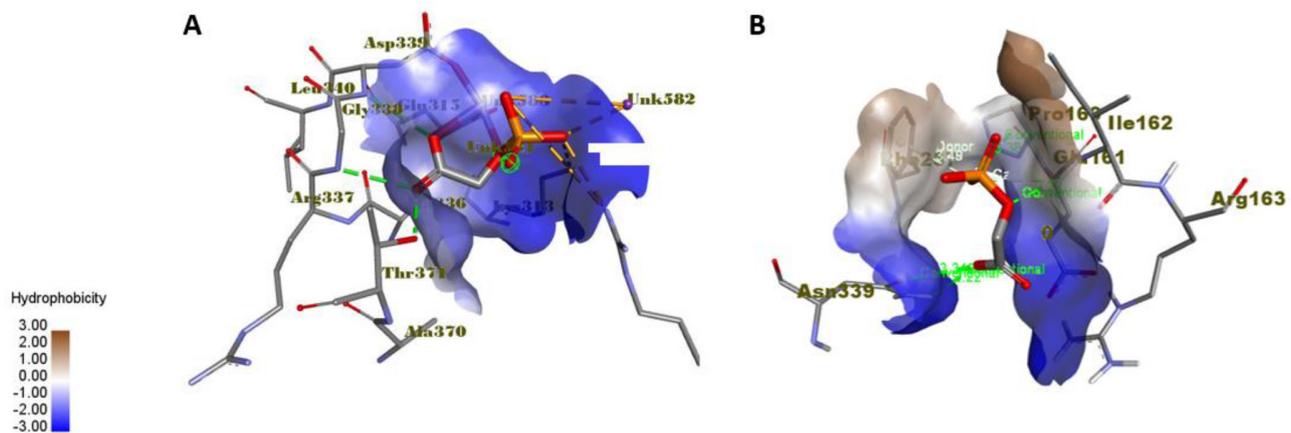


Fig. 16 A comparison in hydrophobicity.

A) the normal domain of PKLR B) The mutated domain with "2-phosphoglycolic acid". Hydrophobicity appeared with brown color, while hydrophilicity appeared with blue color. The docked chain A of normal PKLR was mostly hydrophilic, while the docked mutated PKLR substrate complex had hydrophobic regions.

not form regular interactions between the active site and its modelled substrate (2-phosphoglycolic acid). The normal substrate is 2-phosphoenolpyruvic acid. The phosphoglycolic acid is a substrate-model/inhibitor use in the crystallography because ADP cannot be added with PEP because they would react.

H-bonds are shown to control molecular interactions via a donor-acceptor pairing process that decreases water competition. Moreover, each charge within the protein will communicate with the surrounding solvent. When there are more charges present, those charges interact with one another, but the solvent dampens the intensity of those interactions. There is a substantial impact of charged particle interactions with solvents and solvent filtration of charge-charge interactions on the electrostatic energy of proteins. All these protein properties exhibited significant change between normal and mutated PKLR.

Additionally, according to the 3D modeling of the mutant protein by Rehman and colleagues, with the evaluating of protein-ligand interaction showed that the mutant protein

established new connections through Arg216 and Glu347 and lost its regular interactions with phosphoenolpyruvate, which causes poor or nonexistent dephosphorylation of phosphoenolpyruvate that leads to deficiency in glycolysis's ability to produce energy.¹⁹

Furthermore, we have used several in-silico prediction tools available online to predict pathogenicity or damaging effect of the amino acid substitution on the protein function. The substitution at position 339 from aspartic acid (D) to asparagine (N) is predicted to affect protein function with a score of 0.00 (predicted damaging if the score is equal or less than 0.05), and median sequence conservation of 2.99 (Prediction-sequence diversity is measured using this metric) by using SIFT (sorts intolerant from tolerant), which is a sequence homology-based tool used to estimate the potential phenotypic impact of a protein-level amino acid change. Protein function and evolutionary history are assumed to be linked in SIFT. Alignments of protein families should show conservation at functionally relevant locations while showing diversity at less critical places.³⁴

SNAP² (Screening for Non-Acceptable Polymorphisms) had predicted an effect on protein function as well with a score of 84 and expected accuracy of 91%. SNAP uses a neural network to anticipate how a SNP may influence protein's function. When available, SNAP makes use of evolutionary data to determine which residues within a given sequence family are conserved, as well as other features of the protein's structure and annotations.³⁵

Additionally, another in-silico tool was used to predict the effect of the amino acid substitution, PhD-SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms), which is a machine learning method that employs a supervised training algorithms to determine if a given SNP is connected with disease by analyzing the protein sequence, also resulted in predicting a disease-related polymorphism in corresponding to the change in the amino acid sequence D339N.³⁶ Also, Mutation Assessor, which is based on the evolutionary conservation of the damaged amino acid in protein homologs, provides an approximate assessment of the chance that the mutation has a phenotypic effect at the organism level. This web-based tool has predicted a high functional impact of this missense mutation on the resulted protein.³⁷

PANTHER tool calculates the probability that a particular SNP will lead to a change in the protein's function and presents the outcome as a percentage. It does this by computing a score known as subPSEC, which stands for substitution position-specific evolutionary conservation, based on an alignment of evolutionarily related proteins using a hidden Markov model.³⁸ PANTHER predicted the variation as probably damaging with a probability of deleterious effect score of 0.95.

Moreover, MutPred2 was also used in pathogenicity prediction, which resulted in a score of 0.92 (equal or more than 0.5 is pathogenic).³⁹

Furthermore, the mutation was also categorized as "Probably damaging" by using another in-silico software tool, PolyPhen-2, which is a Bayesian classifier that determines the likelihood that a given non-synonymous SNP is deleterious. The Location-Specific Independent Count (PSIC) score, which ranges from 0 to 1, is based on the evolutionary conservation of a protein sequence in the MSA (multiple sequence alignment) and the negative effect on the protein structure.⁴⁰

Conclusion

With the dramatic increase of the number of molecular mutations, and clinical misdiagnosis, NGS technologies have led to increased knowledge of rare congenital conditions. While none of the proband's family members were clinically affected or had the disease, it was present in the proband in a homozygous state. Our data collectively suggest that c.1015G>A missense variation of the PKLR gene may decrease or abolish PK enzymatic activity, resulting in PK deficiency in the affected individuals. We have submitted this variation to the ClinVar database as pathogenic with association to pyruvate kinase deficiency of the red blood cells with the accession ID # SCV003845976 (processed and released).

To further expand our knowledge of this variation and show its pathogenic character, functional consequences of the variant should be investigated experimentally such as PK enzymatic assays, biomarkers, and expression analysis should be performed in future work. ■

References

- Bianchi, P. and E. Fermo, *Molecular heterogeneity of pyruvate kinase deficiency*. Haematologica, 2020. 105(9): p. 2218–2228.
- Selwyn, J.G. and J.V. Dacie, *Autohemolysis and Other Changes Resulting from the Incubation in Vitro of Red Cells from Patients with Congenital Hemolytic Anemia*. Blood, 1954. 9(5): p. 414–438.
- Valentine, W.N., *A specific erythrocyte glycolytic enzymedefect (pyruvate kinase) in three subjects with congenital nonspherocytic hemolytic anemia*. Trans Ass Amer Physicians, 1961. 74: p. 100–110.
- Tanaka, K.R., W.N. Valentine, and S. Miwa, *Pyruvate kinase (PK) deficiency hereditary nonspherocytic hemolytic anemia*. Blood, 1962. 19: p. 267–95.
- De Gruchy, G.C., et al., *Nonspherocytic congenital hemolytic anemia*. Blood, 1960. 16: p. 1371–97.
- Grace, R.F. and W. Barcellini, *Management of pyruvate kinase deficiency in children and adults*. Blood, 2020. 136(11): p. 1241–1249.
- Al-Samkari, H., et al., *The variable manifestations of disease in pyruvate kinase deficiency and their management*. Haematologica, 2020. 105(9): p. 2229–2239.
- Roy, M.K., et al., *Red Blood Cell Metabolism in Pyruvate Kinase Deficient Patients*. Frontiers in Physiology, 2021. 12.
- Porter, M.L. and B.L. Dennis, *Hyperbilirubinemia in the term newborn*. Am Fam Physician, 2002. 65(4): p. 599–606.
- Grace, R.F., et al., *Clinical spectrum of pyruvate kinase deficiency: Data from the Pyruvate Kinase Deficiency Natural History Study*. Blood, 2018. 131(20): p. 2183–2192.
- Secrest, M.H., et al., *Prevalence of pyruvate kinase deficiency: A systematic literature review*. Eur J Haematol, 2020. 105(2): p. 173–184.
- Canu, G., et al., *Red blood cell PK deficiency: An update of PK-LR gene mutation database*. Blood Cells Mol Dis, 2016. 57: p. 100–9.
- Bianchi, et al., *Addressing the diagnostic gaps in pyruvate kinase deficiency: Consensus recommendations on the diagnosis of pyruvate kinase deficiency*. American Journal of Hematology, 2019. 94(1): p. 149–161.
- Agarwal, A.M., et al., *Clinical utility of next-generation sequencing in the diagnosis of hereditary haemolytic anaemias*. Br J Haematol, 2016. 174(5): p. 806–14.
- Roy, N.B., et al., *A novel 33-Gene targeted resequencing panel provides accurate, clinical-grade diagnosis and improves patient management for rare inherited anaemias*. Br J Haematol, 2016. 175(2): p. 318–330.
- Wooderchak-Donahue, W.L., et al., *A direct comparison of next generation sequencing enrichment methods using an aortopathy gene panel- clinical diagnostics perspective*. BMC Med Genomics, 2012. 5: p. 50.
- Sun, Y., et al., *Next-generation diagnostics: Gene panel, exome, or whole genome? Hum Mutat, 2015. 36(6): p. 648–55.*
- Kim, Y., J. Park, and M. Kim, *Diagnostic approaches for inherited hemolytic anemia in the genetic era*. Blood Res, 2017. 52(2): p. 84–94.
- Rehman, A.U., et al., *A novel homozygous missense variant p.D339N in the PKLR gene correlates with pyruvate kinase deficiency in a Pakistani family: A case report*. J Med Case Rep, 2022. 16(1): p. 66.
- Chen, M., J. Zhang, and J.L. Manley, *Turning on a fuel switch of cancer: HnRNP proteins regulate alternative splicing of pyruvate kinase mRNA*. Cancer Res, 2010. 70(22): p. 8977–80.
- P., B., et al., *Addressing the diagnostic gaps in pyruvate kinase deficiency: Consensus recommendations on the diagnosis of pyruvate kinase deficiency*. Am J Hematol, 2019. 94(1): p. 149–161.
- Schormann, N., et al., *An overview of structure, function, and regulation of pyruvate kinases*. Protein Sci, 2019. 28(10): p. 1771–1784.
- Valentini, G., et al., *Structure and function of human erythrocyte pyruvate kinase. Molecular basis of nonspherocytic hemolytic anemia*. J Biol Chem, 2002. 277(26): p. 23807–14.
- Zanella, A., et al., *Pyruvate kinase deficiency: The genotype-phenotype association*. Blood Reviews, 2007. 21(4): p. 217–231.
- Murakami, K. and M. Yoshino, *Zinc inhibition of pyruvate kinase of M-type isozyme*. Biometals, 2017. 30(3): p. 335–340.

26. Wallace, A.C., R.A. Laskowski, and J.M. Thornton, *LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions*. Protein Eng, 1995. 8(2): p. 127–34.
27. Ben, A., et al., *ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins*. Protein Sci, 2020. 29(1): p. 258–267.
28. Lu, S., et al., *CDD/SPARCLE: The conserved domain database in 2020*. Nucleic Acids Res, 2020. 48(D1): p. D265–d268.
29. Zanella, A., et al., *Molecular Characterization of PK-LR Gene in Pyruvate Kinase-Deficient Italian Patients*. Blood, 1997. 89(10): p. 3847–3852.
30. Zanella, A., et al., *Red cell pyruvate kinase deficiency: Molecular and clinical aspects*. Br J Haematol, 2005. 130(1): p. 11–25.
31. Valentini, G., et al., *The allosteric regulation of pyruvate kinase*. J Biol Chem, 2000. 275(24): p. 18145–52.
32. Mattevi, A., et al., *Crystal structure of Escherichia coli pyruvate kinase type I: molecular basis of the allosteric transition*. Structure, 1995. 3(7): p. 729–41.
33. Mohamed, D.S., et al., *Sesame oil ameliorates valproic acid-induced hepatotoxicity in mice: Integrated in vivo-in silico study*. J Biomol Struct Dyn, 2022: p. 1–21.
34. Ng, P.C. and S. Henikoff, *SIFT: Predicting amino acid changes that affect protein function*. Nucleic Acids Res, 2003. 31(13): p. 3812–4.
35. Bromberg, Y. and B. Rost, *SNAP: Predict effect of non-synonymous polymorphisms on function*. Nucleic Acids Res, 2007. 35(11): p. 3823–35.
36. Capriotti, E. and P. Fariselli, *PhD-SNPg: A webserver and lightweight tool for scoring single nucleotide variants*. Nucleic Acids Res, 2017. 45(W1): p. W247–W252.
37. Reva, B., Y. Antipin, and C. Sander, *Predicting the functional impact of protein mutations: Application to cancer genomics*. Nucleic Acids Res, 2011. 39(17): p. e118.
38. Thomas, P.D., et al., *PANTHER: A library of protein families and subfamilies indexed by function*. Genome Res, 2003. 13(9): p. 2129–41.
39. Pejaver, V., et al., *Inferring the molecular and phenotypic impact of amino acid variants with MutPred2*. Nat Commun, 2020. 11(1): p. 5918.
40. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations*. Nat Methods, 2010. 7(4): p. 248–9.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License which allows users to read, copy, distribute and make derivative works for non-commercial purposes from the material, as long as the author of the original work is cited properly.